COMPLEX SYSTEM MODELING AND SIMULATION
ISSN 2096-9929 04/06 pp 308−321
Volume 1, Number 4, December 2021
DOI: 10.23919/CSMS.2021.0026

Hybrid Deep Learning Model for Short-Term Wind Speed Forecasting Based on Time Series Decomposition and Gated Recurrent Unit

Changtong Wang, Zhaohua Liu*, Hualiang Wei, Lei Chen, and Hongqiang Zhang

Abstract: Accurate wind speed prediction has been becoming an indispensable technology in system security, wind energy utilization, and power grid dispatching in recent years. However, it is an arduous task to predict wind speed due to its variable and random characteristics. For the objective to enhance the performance of forecasting short-term wind speed, this work puts forward a hybrid deep learning model mixing time series decomposition algorithm and gated recurrent unit (GRU). The time series decomposition algorithm combines the following two parts: (1) the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), and (2) wavelet packet decomposition (WPD). Firstly, the normalized wind speed time series (WSTS) are handled by CEEMDAN to gain pure fixed-frequency components and a residual signal. The WPD algorithm conducts the second-order decomposition to the first component that contains complex and high frequency signal of raw WSTS. Finally, GRU networks are established for all the relevant components of the signals, and the predicted wind speeds are obtained by superimposing the prediction of each component. Results from two case studies, adopting wind data from laboratory and wind farm, respectively, suggest that the related trend of the WSTS can be separated effectively by the proposed time series decomposition algorithm, and the accuracy of short-time wind speed prediction can be heightened significantly mixing the time series decomposition algorithm and GRU networks.

Key words: deep learning; complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN); gated recurrent unit (GRU); short term; wavelet packet decomposition; wind speed prediction

1 Introduction

Wind power has been vigorously developed because it is clean and renewable. A great number of wind plants have been constructed, and offshore wind power has also received plenty of attention. An arduous challenge when using wind energy is to maintain the wind power grid stability and security because of its randomness, volatility, and variability[1]. Therefore, precisely predicting short-term wind speed (e.g., from 0.5 hour to 3 hour ahead prediction) in a wind farm is essential for efficient power grid dispatching, system security, and optimal operation.

Due to its highly variable and random characteristics, wind speed time series (WSTS) are usually modeled using complex nonlinear quantitative analysis techniques. The existing mainstream wind speed prediction techniques generally include the following kinds: (1) physical methods, (2) linear regression methods, (3) nonlinear artificial intelligence methods, and (4) hybrid methods that combine two or more different approaches. Physical methods, such as numeric weather prediction (NWP), make use of

• Changtong Wang, Zhaohua Liu, Lei Chen, and Hongqiang Zhang are with the School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China. E-mail: ctwang@mail.hnust.edu.cn; zhaohualiu2009@hotmail.com; chenlei@hnust.edu.cn; hongqiangzhang@hnust.edu.cn.
• Hualiang Wei is with the Department of Automatic Control and Systems Engineering, the University of Sheffield, Sheffield, S1 3JD, UK. E-mail: w.hualiang@sheffield.ac.uk.
* To whom correspondence should be addressed.
Manuscript received: 2021-08-24; revised: 2021-10-11; accepted: 2021-11-16

meteorological knowledge (humidity and temperature, etc.), and are usually more suitable for medium to long-term prediction[2]. Linear regression methods, as a low-cost technique, can generate model based on historical data. The primary linear regression models include autoregression (AR) model and autoregression moving average (ARMA) model[3]. Karakuş et al.[4] employed polynomial AR model to forecast wing speed 24 hours in advance utilizing hourly WSTS. Linear methods are usually easy to implement but normally they cannot sufficiently capture the nonlinearities in wind speed signals.

Artificial intelligence methods, developed for approximating nonlinear behavior involved in data, can be employed to handle random and nonstationary WSTS. As the typical example of artificial intelligence, artificial neural network (ANN) has extensively been applied in forecasting wind speed. For instance, in Ref. [5], the parameters of wavelet neural networks (WNNs) were optimized adopting evolutionary algorithms for forecasting short-term WSTS. ANN is deemed to be a satisfactory tool with a high prediction performance. However, such a method may suffer from some drawbacks such as slow learning, overfitting, and difficulty in deciding hyper-parameters[6]. Moreover, a simple ANN structure may not be able to sufficiently capture complex patterns from the random WSTS and therefore is weak in generalization[7]. Except ANN, support vector machine (SVM) can also perform time series prediction effectively. In Ref. [8], a method combined with Markov model and SVM was presented, aiming at solving the dynamic problems of wind ramps in wing power prediction. Based on the measured data, this approach integrated Markov model into an enhanced SVM and obtained a high prediction accuracy. It is useful to note that in general SVM is more suitable for classification than regression problems. In addition, in Ref. [9], a Gaussian process (GP) probabilistic method based on temporally local "moving window" approach was proposed for dramatically reducing the prediction error. Compared with traditional artificial intelligence technology, deep learning can better capture the intrinsic nonlinear relationship of the process of interest by exploiting information from big data[10]. Deep generative neural network[11] and long short-term memory (LSTM) network, as reliable deep learning approaches, are widely employed to conquer the time dependent problem[12]. In Ref. [13], an LSTM model combined with fuzzy-rough set theory was designed, and achieved good wind speed prediction results. As a simplification of LSTM, gated recurrent unit (GRU) has similar performance to LSTM. With fewer parameters, GRU is faster to train and less prone to overfitting[14]. In Ref. [15], combined with variational mode decomposition (VMD), GRU networks were integrated to forecast wind speed interval and gained a higher prediction interval coverage probability.

WSTS is complex because of its random and fluctuant characteristics. Any single method aforementioned may suffer some inherent shortages when directly applied to WSTS. Therefore, the development of hybrid methods to achieve better prediction accuracy has attracted extensive attention[16]. Such a method combines two or more good algorithms into a single one to strengthen the analysis ability[17]. For achieving higher prediction accuracy, Shi et al.[18] devised a available mixed artificial intelligence technique by discussing wind speed distribution and optimizing model weights to predict wind power. Khodayar et al.[19] used the stack versions of autoencoder (AE) and denoising AE in the deep neural network (DNN) to dispose of the uncertainty of WSTS. However, deep learning approaches (e.g., DNN) cannot easily capture the hidden complex pattern of WSTS because the noise exists in the input data which include quite rich frequency components. Therefore, the idea of signal decomposition and reconstruction is often utilized to facilitate handling complex WSTS; such an approach can help reduce the complexity of dealing with WSTS using whatever data modeling methods, and thus can help improve prediction performance. Commonly used signal decomposition and reconstruction methods include empirical mode decomposition (EMD)[20, 21], wavelet transform (WT)[22], and wavelet packet decomposition (WPD)[23, 24]. The eventual forecast values can be gained by superimposing the prediction results of all the decomposed components. In Ref. [25], a deep learning architecture considering LSTM networks and VMD algorithm was designed to boost the ability of multistep wind power prediction, where VMD was built to transform raw series into several sub-modes. Zheng et al.[26] employed WPD algorithm to reduce the complexity of WSTS and then designed a particle swarm optimization training algorithm with disturbance to calculate the weights of improved Elman neural network, and achieved good effect. Based on

meteorological information, Haque et al.[27] built a mixed algorithm with WT and fuzzy adaptive resonance theory mapping (ARTMAP) network to predict wind power, where WT was used to extract potential features and improve the ability of the fuzzy ARTMAP network. In Ref. [28], an effective nonstationary WSTS prediction technique was studied, where EMD and least square support vector machine were mixed to predict one-month ahead wind speed. As mentioned in Refs. [29, 30], EMD has a good adaptive property and can decompose WSTS into several intrinsic mode functions (IMFs) and a residue with different frequencies. Nonetheless, incomplete decomposition causes each IMF to mix more than two frequencies (called mode mixing problem)[31]. For the purpose of overcoming the defect, many methods have been proposed[32−34], e.g., the ensemble EMD (EEMD) and the complete EEMD with adaptive noise (CEEMDAN).

Although a large number of prediction methods have been presented to solve the problem of wind speed forecasting, there are still some shortcomings for these existing methods: (1) Because the change of wind speed is temporal, it is difficult for general machine learning and deep learning algorithms to capture the latent information of WSTS. (2) Some existing prediction methods focus on the improvement of machine learning and deep learning algorithms themselves, and do not adequately process WSTS, which is not easy to mine the features hidden in complex and fluctuating WSTS. (3) Some signal decomposition techniques are applied to process WSTS in other methods, but these techniques cannot adequately capture the potential trends of complex and fluctuating WSTS due to their own defects, such as mode mixing problem in EMD technique. Another disadvantage of EMD is that the frequency and amplitude of the first IMF may fluctuate greatly, and this can seriously influence the prediction accuracy.

For the objective to enhance the forecasting performance of short-term wind speed, a hybrid deep learning model is proposed. This model integrates a time series decomposition algorithm and GRU networks, and has the following characteristics: (1) Several IMFs and a residue can be obtained by employing CEEMDAN to decompose the normalized WSTS; (2) WPD is adopted to further decompose the first IMF that contains the complex and high-frequency signal of the raw data; (3) GRU networks are established for all the resulting component signals, including training samples and predicting the outputs; (4) the eventual predicted values are calculated by superimposing the prediction results of all components. The proposed method can handle complex WSTS and help deal with the randomness and fluctuation in WSTS, and therefore improves the prediction accuracy. The major contributions of this work are as follows:

(1) An effective time series decomposition algorithm is proposed. This algorithm can capture the complex patterns of WSTS by using two decomposition algorithms, CEEMDAN and WPD. The proposed algorithm is useful to dispose of the randomness and fluctuation in WSTS.

(2) For predicting short-term wind speed, a mixed model is designed by fusing GRU networks with time series decomposition. As a deep learning algorithm, GRU can effectively capture the nonlinear fluctuation of WSTS, therefore can significantly enhance the forecasting performance aiming at short-term wind speed. A comparison is presented for the illustration of the effectiveness of the proposed method and other seven methods.

The rest sections are arranged as follows. The proposed hybrid deep learning approach is introduced in Section 2. The results compared with other methods through two cases are described in Section 3. This main work is summarized in Section 4.

2 Proposed Hybrid Deep Learning Model Based on Time Series Decomposition and GRU for Short-Term Wind Speed Forecasting

2.1 Architecture of the proposed deep learning method

For the objective of heightening the forecasting accuracy, the design of the proposed hybrid deep learning method comprises the following main steps: (1) the design of the time series decomposition algorithm, and (2) the design of GRU networks based prediction model. The design of the time series decomposition algorithm aims to capture the complex patterns of WSTS. The GRU networks based prediction model is designed to handle the time dependent problem of WSTS based on the its gate mechanism. The eventual predicted values can be calculated by superimposing the prediction results for the individual components resulted from the decomposition. For explaining the model, Fig. 1 gives the structure, and the

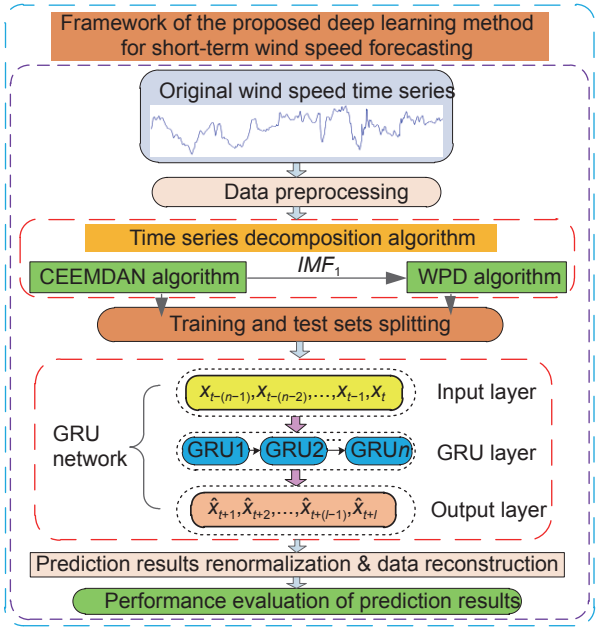

Fig. 1 Structure of the designed deep learning model.

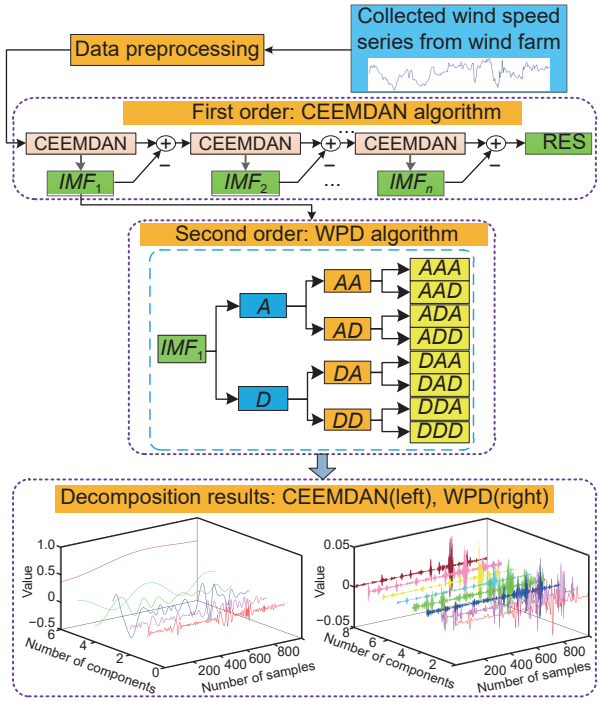

Fig. 2 Framework of the time series decomposition algorithm.

design is described in detail in the following sections.

2.2 Time series decomposition algorithm for WSTS

The original WSTS may be highly random and nonstationary in nature, so reducing the complexity of WSTS is often a useful and important initial step towards achieving good wind speed forecasting. Consequently, the paper designs a novel time series decomposition algorithm. It integrates two algorithms, i.e., the CEEMDAN and WPD algorithms, which enable capturing the complex patterns from WSTS. Figure 2 shows the framework of the time series decomposition algorithm.

The CEEMDAN algorithm is developed from EEMD to adaptively decompose the nonstationary and nonlinear signals. EEMD can average the modes obtained by EMD by adding Gaussian white noise to obtain IMFs, but these IMFs may not meet the constrains and cannot reconstruct original signal accurately due to the existence of noise. Additionally, the computational cost is larger. Compared with EEMD, CEEMDAN adds adaptive Gaussian white noise in each stage of decomposition to eliminate the false IMFs. It can precisely reconstruct the original signal and obtain purer IMFs. Furthermore, its computational cost is reduced due to the less use of sifting iterations than EEMD[35]. Here, the CEEMDAN algorithm is adopted in the first-order decomposition process to handle the normalized WSTS, and the detailed steps are given in the following procedure:

(1) Random Gaussian white noise $x_i(t)$ is blended into normalized WSTS $s(t)$ and obtain noise-added WSTS: $S_i(t) = s(t) + \omega_0 x_i(t)$, where $i = 1, 2, \ldots, I$, and ω_0 denotes the noise coefficient;

(2) The first component $IMF_{1,i}$ can be calculated by decomposing $S_i(t)$ using EMD, and then take an average

$$IMF_1 = \frac{1}{I}\sum_{i=1}^{I} IMF_{1,i} \tag{1}$$

(3) The first residue is defined as $r_1(t) = s(t) - \overline{IMF_1}$;

(4) Decompose $r_1(t) + \omega_1 E_1(x_i(t))$ to obtain the second IMF

$$IMF_2 = \frac{1}{I}\sum_{i=1}^{I} E_1(r_1(t) + \omega_1 E_1(x_i(t))) \tag{2}$$

where ω_1 is the noise coefficient and $E_j(\cdot)$ denotes that achieving the j-th IMF using EMD;

(5) Iteratively run steps (3) and (4) until there are no more than two extrema in the residue;

(6) The function expression of normalized WSTS decomposed by CEEMDAN is written as

$$s(t) = \sum_{i=1}^{I} IMF_i + RES \tag{3}$$

where IMF_i is the i-th IMF decomposed by

CEEMDAN, and *RES* is the residue of the decomposition process.

The WPD algorithm can perform multiresolution analysis of a signal with WT. As investigated in Ref. [36], WT is the inheritance and development of the Fourier transform. It has a window whose area is fixed and its shape changes with frequency. WPD has been applied to WSTS forecasting[37]. WPD is different from WT in that WPD decomposes both the appropriate and detailed components, while WT only decomposes the appropriate components. Note that the IMF_1 obtained by CEEDMAN consists of the complex and high-frequency component signals of the normalized WSTS. So WPD algorithm is used to further dispose of the IMF_1.

Let h_{0m} and h_{1m} be filter coefficients in multiresolution analysis. The orthogonal scale function $\alpha(\tau)$ and wavelet function $\beta(\tau)$ are as below:

$$\alpha(\tau) = \sqrt{2}\sum_m h_{om}\phi(2\tau - m) \tag{4}$$

$$\beta(\tau) = \sqrt{2}\sum_m h_{1m}\phi(2\tau - m) \tag{5}$$

For convenience, the following recurrence relations are often considered:

$$\delta_{2\upsilon}(\tau) = \sqrt{2}\sum_{m\in Z} h_{0m}w_\upsilon(2\tau - m) \tag{6}$$

$$\delta_{2\upsilon+1}(\tau) = \sqrt{2}\sum_{m\in Z} h_{1m}w_\upsilon(2\tau - m) \tag{7}$$

where $\delta_0(\tau) = \alpha(\tau)$, and $\delta_1(\tau) = \beta(\tau)$.

Now assume that $f(\tau)$ is a signal (e.g., WSTS) to be processed, and $p_j^i(\tau)$ is the i-th wavelet packet on the j-th layer, which is called the wavelet packet coefficient. R and Q correspond the low and high pass wavelet decomposition filters, respectively. The fast algorithm of the dyadic wavelet decomposition is as follows:

$$\begin{cases} p_0^1(\tau) = f(\tau), \\ p_j^{2i-1}(\tau) = \sum_m R(m-2\tau)p_{j-1}^i(\tau), \\ p_j^{2i}(\tau) = \sum_m Q(m-2\tau)p_{j-1}^i(\tau) \end{cases} \tag{8}$$

Its reconstruction algorithm is

$$\begin{aligned} p_j^i(\tau) =& 2\sum_m r(\tau - 2m)p_{j+1}^{2i-1}(\tau) + \\ & 2\sum_m q(\tau - 2m)p_{j+1}^{2i}(\tau) \end{aligned} \tag{9}$$

where $j = F-1, F-2, \ldots, 0$, F denotes the decomposition level, and $i = 2^j, 2^{j-1}, \ldots, 1$. r and q are wavelet decomposition filters.

The part of WPD algorithm shown in Fig. 2 is a three-level decomposition process that decomposes IMF_1 into eight groups of wavelet packet coefficients providing more information than IMF_1 alone, where A and D represent the appropriate and detailed components, respectively. Figure 2 also shows a series of IMFs decomposed by CEEMDAN algorithm. Therefore, through the time series decomposition algorithm, the inherent features of WSTS are decomposed to a batch of components, which are easier to manage and more useful for the prediction. Finally, all the components are reorganized into training and test sets, respectively.

2.3 GRU-based prediction model

Assuming that the input data are $(x_{t-(n-1)}, x_{t-(n-2)}, \ldots, x_t)$, and prediction outputs are $(\hat{x}_{t+1}, \hat{x}_{t+2}, \ldots, \hat{x}_{t+l})$, the prediction task is to predict next l points using previous n points at the moment t. When $l = 1$, it is called one-step prediction; when $l > 1$, it is called multistep prediction. Figure 3 gives the GRU structure and the designed GRU-based prediction model. GRU, in structure is similar to LSTM[38] but involves fewer parameters. In GRU, there are only two gates, namely, only update gate $u(t)$ and reset gate $v(t)$. Therefore, it possesses a more user-friendly form and speedier operating rate than LSTM. The main function of update gate is to determine how many previous time point states are sent to the current moment, while the reset gate is to determine how many previous time point states are sent to the candidate set $\tilde{o}_t$ at the current moment. Let W_u, W_v, and W_o be the weights of update gate, reset gate, and candidate output, respectively. The feedforward propagation process is described following the mathematics below:

$$u_t = \sigma(W_u \cdot [o_{t-1}, x_t]) \tag{10}$$

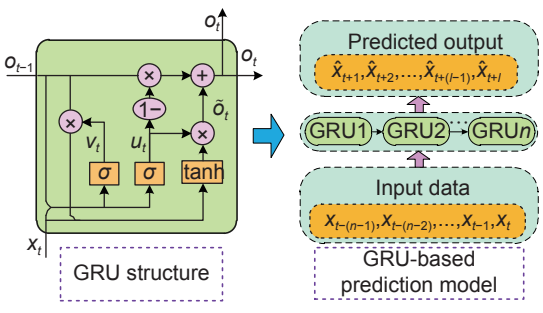

Fig. 3 GRU structure (left) and the proposed GRU-based prediction model (right).

$$v_t = \sigma(W_v \cdot [o_{t-1}, x_t]) \tag{11}$$

$$\tilde{o}_t = \tanh(W_{\tilde{o}} \cdot [v_t \odot o_{t-1}, x_t]) \tag{12}$$

$$h(t) = (1 - u_t) \odot o_{t-1} + u_t \odot \tilde{o}_t \tag{13}$$

where o_{t-1} is the output at the moment $t-1$. x_t is the input at the moment t. σ and tanh are the activation functions.

2.4 Reconstruction and renormalization of prediction results

The prediction results of all the components obtained by the two-stage decomposition algorithm need to be reconstructed to obtain the final results. This is achieved through the following summation:

$$S' = \sum_{i=2}^{n} IMF'_i + RES' + \sum_{j=0}^{2^J-1} WPD'_j \tag{14}$$

where S' is the predicted results of normalized WSTS, IMF'_i is the predicted values of the i-th IMF, RES' is the predicted results of the residue, and WPD'_j represents the predicted results of the j-th group of wavelet packet coefficient. Finally, the final prediction results are obtained by renormalization of S'.

2.5 Evaluation metrics

With a view to assessing the good capability of the model, two commonly used metrics are adopted, that is, normalized root mean square error (*NRMSE*) and normalized mean absolute percentage error (*NMAPE*)[39]:

$$NRMSE = \sqrt{\frac{1}{T}\sum_{\zeta=1}^{T}\left(\frac{\eta_\zeta - \hat{\eta}_\zeta}{\eta_{\max} - \eta_{\min}}\right)^2} \times 100\% \tag{15}$$

$$NMAPE = \frac{1}{T}\sum_{\zeta=1}^{T}\frac{|\eta_\zeta - \hat{\eta}_\zeta|}{\eta_{\max} - \eta_{\min}} \times 100\% \tag{16}$$

where η_ζ and $\hat{\eta}_\zeta$ are the measured and predictive value of WSTS, respectively. $\eta_{\min}$ and $\eta_{\max}$ are the minimum and maximum of the measured values, respectively. T is the length of WSTS.

2.6 Flow of the prediction model

Aiming at forecasting short-time wind speed, the following procedure is presented based on the proposed model:

Step 1: Collect WSTS from wind farm and normalize them into [0, 1];

Step 2: Apply the proposed time series decomposition algorithm to decompose normalized WSTS, and a batch of components will be acquired. Each component is split into training and test sets;

Step 3: Build GRU-based prediction model for each component, and use the training sets to train each model based on the root mean square prop (RMSProp) optimizer. The mean square error (MSE) is selected as the loss function. Then, the prediction accuracy is verified in test sets.

Step 4: The predicted wind speeds are computed by superimposing the prediction results of all components, and then the evaluation metrics are computed according to Eqs. (15) and (16).

More details about the implementation are given in Algorithm 1.

3 Simulation Results and Discussions

With a view to assessing the good capability of the designed forecasting method, the cases with the prediction horizon of 10 min, 30 min, 1 h, and 2 h are conducted on two wind speed datasets. The simulation is performed on the Anaconda3 (64-bit) environment, where the basic features of the computer are as follows: an Inter (R) Core (TM) i5-9400 F, 2.90 GHz CPU and 8.00 GB RAM.

3.1 Case I: Simulation verification in an offshore wind farm

3.1.1 Data description

In Case I, the wind speed dataset was collected from an offshore wind farm[40]. In this wind farm, the wind speed data at site 10 002 were chosen for this case simulation, and the site was located at the latitude 38.393 31 and longitude −74.936 54. The sample interval of this dataset is 10 min. For verifying the proposed model adequately, the data for four seasons were considered in this case. Specifically, the data for the first week of each season in 2004 were selected, forming four datasets to test the performance of the proposed method in spring, summer, autumn, and winter, respectively. For each dataset, the data for the first six days containing 864 sample points were used as the training set, and the data for the last day were regarded as the test set. Considering the four forecast horizons, a total of 16 sets of simulations were conducted.

3.1.2 Implementation details and parameter settings

In this case simulation, the model parameters of the proposed method need to be set. For the above corresponding forecast horizons, the number of output nodes for GRU-based prediction model is 1, 3, 6, and

Algorithm 1 Flow of the prediction model

Input: Original WSTS $s(t)$.

Output: Prediction results $s'(t)$.

1: Function *Data_Preprocessing*($s(t)$);

1-1: $Ns(t) \leftarrow \frac{s(t)-s(t)_{\min}}{s(t)_{\max}-s(t)_{\min}}$;

1-2: return $Ns(t)$.

2: Function *Decomposition*($Ns(t)$);

2-1: for i=1 to I **do**

$IMF_i \leftarrow CEEMDAN(Ns(t))$;

$Ns(t) \leftarrow Ns(t) - IMF_i$;

end (for loop)

$RES \leftarrow Ns(t)$;

2-2: $imf_0^1 \leftarrow IMF_1$;

for j=1 to 2 **do**

for i=1 to j **do**

$imf_j^{2i-1} \leftarrow \sum_k H(k-2t) imf_{j-1}^i$;

$imf_j^{2i}(t) \leftarrow \sum_k G(k-2t) imf_{j-1}^i$;

end (for loop)

end (for loop)

for j=0 to 2^F-1 **do**

Reconstruct each wavelet packet coefficient by using Eq. (9) while other coefficients are set as zero;

end (for loop)

2-3: return all the components obtained by CEEMDAN and WPD (except IMF_1);

3: Function *Prediction*(*all the components*);

3-1: for *all the components* **do**

$GRU\_input \leftarrow (x_{t-(n-1)}, x_{t-(n-2)}, \ldots, x_t)$;

$GRU\_output \leftarrow (x_{t+1}, x_{t+2}, \ldots, x_l)$;

$training\_set, test\_set \leftarrow each\ component$;

for *epoch*=1 to *epochs* **do**

GRU() ← train the GRU model with training_set;

end (for loop)

$predicted \leftarrow GRU\ (test\_set)$;

$S' \leftarrow S' + predicted$;

end (for loop)

3-2: $s'(t) \leftarrow$ renormalize S'.

3-3: return prediction results $s'(t)$.

12, respectively. Considering the computational cost and prediction efficiency, the hidden layer number was preset as one, and the number of hidden layer neurons is chosen from 2 to 40 to obtain the optimal results for each component. In addition, as the decomposition level for WPD algorithm increases, the number of wavelet packet coefficients will become larger and the time cost will increase. Therefore, the two-level WPD algorithm was adopted in this study, and four groups of wavelet packet coefficients (*AA*, *AD*, *DA*, and *DD*) are obtained.

Additionally, the prediction performance was compared with other seven methods, namely, back propagation neural network (BP), LSTM, GRU, convolutional neural networks based LSTM (CNN-LSTM), EMD-based GRU (EMD-GRU), WPD-based GRU (WPD-GRU), and CEEMDAN-based GRU (CEEMDAN-GRU).

For realizing the equitable comparison, the optimal parameters are selected for all the comparison methods in each group of simulation. In these methods, the number of hidden layers is set to 1, and the number of hidden layer neurons is chosen from 2 to 40 following different datasets. In addition, the optimizer of these methods based on LSTM and GRU models is set to RMSProp whose learning rate is 0.001. For BP method, the number of hidden layer neurons is set to 6, and the optimizer is stochastic gradient descent (SGD) algorithm whose learning rate is 0.1. For CNN-LSTM method, the one-dimensional CNN is chosen, and the output data of the CNN are used as the input data of LSTM. Correspondingly, the number of the filters, kernels, and strides in CNN is set to 8, 2, and 1, respectively. For the EMD-GRU and CEEMDAN-GRU methods, the WSTS are adaptively decomposed following the same method as the proposed model. For WPD-GRU method, the number of the decomposition levels of WPD algorithm is set to 2. For these decomposition-based methods, the optimal model parameters of GRU are determined for each component obtained from the decomposition techniques.

3.1.3 Result analysis in Case I

Aiming at wind speed forecasting, Tables 1 and 2 show the *NRMSE* and *NMAPE* values of all the experimental methods in four seasons (the forecast horizon is from 10 min to 2 h). Bold font is used to highlight the best prediction performance. The histograms of two metrics are displayed in Fig. 4.

In terms of *NRMSE* and *NMAPE*, the proposed model has better-performing capability than other seven techniques in all the seasons. Taking Fig. 4a (spring) as an example, in comparison with the BP method, the *NRMSE* values decrease 48.71% and 55.31% for 10 min forecast horizon (from 4.8671% to 2.4963%) and 2 h forecast horizon (from 22.5947% to 10.0978%), respectively, and the *NMAPE* value decreases 53.77% and 54.64% for 10 min forecast horizon (from 3.7551% to 1.7361%) and 2 h forecast horizon (from 16.7187% to 7.5842%), respectively. As

Table 1 ***NRMSE*** **values of forecasting methods in four seasons in Case I**

Method	*NRMSE* value (%)															
	10 min				30 min				1 h				2 h			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
BP	4.8671	3.8133	2.6768	3.8876	10.3562	6.9453	5.4511	8.6173	16.6296	11.4874	8.3336	14.1404	22.5947	14.6451	12.3007	18.6259
LSTM	4.3898	3.1618	2.4825	3.7688	9.3786	6.6365	5.1206	8.0946	16.1849	11.0527	7.7266	13.8370	23.2575	14.5829	11.5804	17.4295
GRU	4.2172	2.9557	2.3895	3.5603	9.1842	6.5100	4.4829	7.9418	15.6005	11.0326	7.4736	12.9447	21.8145	14.1907	11.4867	17.7750
CNN-LSTM	3.8495	3.4293	2.4481	3.7216	9.1986	6.5121	4.7164	7.8398	14.7261	10.6396	7.2302	11.8735	18.2457	14.1106	11.3983	17.3029
EMD-GRU	3.7293	1.8322	1.9855	6.8295	5.4917	3.2266	2.5465	11.6864	9.0159	5.5553	2.7386	11.7137	11.0420	7.4532	5.7003	11.5743
WPD-GRU	3.8130	2.5739	1.4271	3.1020	8.8071	6.4667	3.1443	7.0214	15.0796	10.0255	7.4531	11.8591	22.8781	14.8975	11.4535	18.6366
CEEMDAN-GRU	3.5004	1.9214	1.3699	2.5331	4.8149	3.4212	2.2759	4.5437	7.9570	5.4629	2.6185	8.1481	10.3293	7.3702	4.3625	10.9419
Proposed method	**2.4963**	**1.4237**	**1.0411**	**2.0777**	**4.2327**	**2.9805**	**2.0406**	**4.0528**	**7.5206**	**5.2478**	**2.6139**	**7.8310**	**10.0978**	**7.3506**	**4.3141**	**10.8894**

Table 2 ***NMAPE*** **values of forecasting methods in four seasons in Case I**

Method	*NMAPE* value (%)															
	10 min				30 min				1 h				2 h			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
BP	3.7551	2.8333	1.9803	2.8575	7.8047	4.9655	4.0536	6.2964	12.4003	8.2214	6.0004	10.7759	16.7187	10.5549	9.2463	13.8531
LSTM	3.3181	2.1445	1.9217	2.8348	6.9866	4.5856	3.8192	5.5780	12.4255	7.4624	5.6688	10.2204	18.3192	10.1924	8.7832	11.7672
GRU	3.2366	2.1053	1.8553	2.6160	6.7523	4.5816	3.4091	5.4161	12.0487	7.5253	5.5689	9.4660	17.5430	10.0521	9.0767	11.8846
CNN-LSTM	2.9451	2.6594	1.9316	2.8951	6.6824	4.4054	3.5949	5.5131	11.3152	6.8293	5.3240	8.1935	13.6383	10.0034	8.9370	12.0310
EMD-GRU	2.9525	1.34558	1.7073	4.2095	4.2064	2.3132	2.0308	7.0771	7.4096	3.4517	2.1590	8.5166	8.8310	4.7407	4.3320	8.9515
WPD-GRU	2.8214	1.9972	1.1310	2.6239	6.1009	4.7249	2.2471	5.0333	10.9180	6.7854	5.2560	8.2447	17.2483	10.6898	8.5163	13.9486
CEEMDAN-GRU	2.6769	1.2304	0.9666	2.0680	3.5799	2.3505	1.7615	3.3091	5.9996	3.8680	2.0023	5.9850	7.6374	5.2512	3.3687	8.6647
Proposed method	**1.7361**	**0.9350**	**0.8708**	**1.7154**	**3.1787**	**2.0417**	**1.6585**	**2.8558**	**5.4167**	**3.7156**	**1.9808**	**5.6074**	**7.5842**	**5.2455**	**3.2882**	**8.5102**

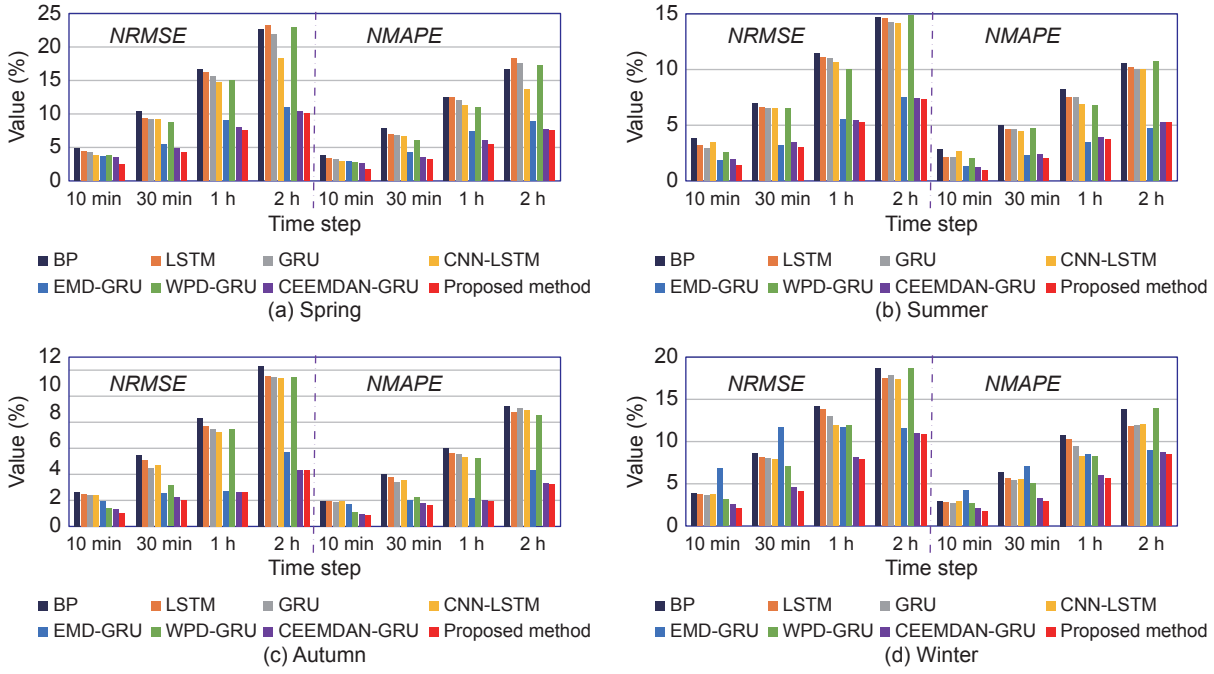

Fig. 4 Histogram of all the metrics in four seasons in Case I.

another example, in Fig. 4c (autumn), compared with the simple GRU method, the *NRMSE* values decrease 56.43% and 62.44% for 10 min forecast horizon (from 2.3895% to 1.0411%) and 2 h forecast horizon (from

11.4867% to 4.3141%), respectively, and the *NMAPE* values decrease 53.06% and 63.77% for 10 min forecast horizon (from 1.8553% to 0.8708%) and 2 h forecast horizon (from 9.0767% to 3.2882%), respectively. The major reasons of this proposed method having higher accuracy may be as follows: (1) GRU network has good time memory ability and is more appropriate to dispose of WSTS than BP. It is obvious from Fig. 4 that the performance of GRU network is better than that of BP method in all the simulation. Moreover, the results suggest that GRU, as a variant of LSTM, can carry out more pinpoint wind speed prognosis than LSTM. (2) The WSTS itself is stochastic, nonlinear, and nonstationary, which is difficult to predict. However, the proposed time series decomposition algorithm can transform the WSTS into some subsequences, which are easier to utilize for forecasting task. In addition, the forecasting capability of all the methods decreases as the forecast horizon increases from 10 min to 2 h, which means that it is much difficult to perform multistep wind speed prediction. It is also seen from Fig. 4 that the prediction results of the proposed method are more accurate than those of other methods in all the seasons. This means that the proposed method has better robustness in processing complex WSTS of different seasons.

In addition, it can also be found that the metrics for the CEEMDAN-GRU method and the proposed model are very close for 2 h ahead wind speed prediction in each season. For example, the *NRMSE* values of the CEEMDNA-GRU and the proposed methods are 7.3702% and 7.3506% in summer, respectively, and are 10.9419% and 10.8894% in winter, respectively. That means that the superiority of WPD algorithm applied in the proposed method may be reduced in pace with the increase of the forecast horizon. More obviously, achieving accurate forecasting is much arduous for 2 h ahead prediction compared with 10 min ahead prediction. From Tables 1 and 2 as well as Fig. 4, it is also obvious that compared with other techniques, the metric values of the EMD-GRU, CEEMDAN-GRU, and the proposed method were relatively low in most cases, and CEEMDAN algorithm outperforms EMD algorithm. For instance, in the case of winter, the performance of EMD-GRU method is very poor than CEEMDAN-GRU method. That is because the serious mode mixing problem occurred in the process of EMD decomposing the WSTS, which greatly reduced the prediction accuracy. For WPD-GRU method, its performance is similar with other methods without considering decomposition algorithm, which may be because that WPD algorithm cannot fully extract the trends of WSTS.

Figure 5 exhibits the prediction curve of the proposed method at 10 min forecast horizon in spring. In Fig. 5 the two lines almost overlap, which means that the proposed model has ability to well forecast the variations of wind speed.

3.2 Case II: Simulation verification in an onshore wind farm

3.2.1 Data description

For much specifically validating its performance, this case considered an onshore wind farm application. The dataset was gathered from site 7856 of a real-world wind farm, which was located latitude 40.244 30 and longitude −85.292 68. In this dataset, four groups of data from four seasons in 2006 were selected, with the sampling interval of 10 min. In each group of data, the data for the first six days were used as the training set, and the data for the last day were considered as the test set. The parameter settings of the proposed method and comparison methods were same with Case I.

3.2.2 Result analysis in Case II

The evaluation metrics of all the methods for four seasons in this case are presented in Tables 3 and 4, and the corresponding histograms are displayed in Fig. 6, from which the proposed model has lowest *NRMSE* and *NMAPE* values. For example, in Fig. 6a (spring), in comparison with the BP method, the *NRMSE* values decrease 71.15% and 65.74% for 10 min forecast horizon (from 4.5200% to 1.3041%) and 2 h forecast horizon (from 12.7084% to 4.3533%), respectively, and the *NMAPE* values decrease 69.76% and 66.48% for 10 min forecast horizon (from 3.4970% to 1.0576%) and 2 h forecast horizon (from 9.9571% to

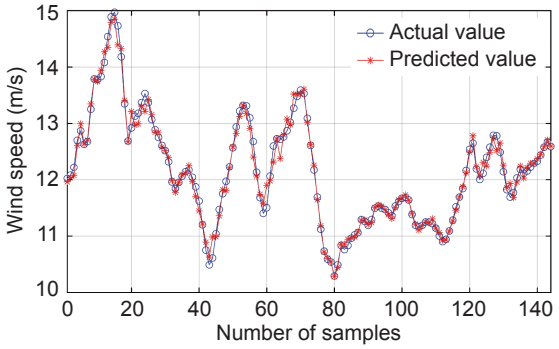

Fig. 5 Prediction curve of the proposed method at 10 min forecast horizon in spring in Case I.

Table 3 ***NRMSE*** **values of forecasting methods in four seasons in Case II**

Method	*NRMSE* value (%)															
	10 min				30 min				1 h				2 h			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
BP	4.5200	11.3828	4.4399	5.4553	5.6881	13.6854	9.5708	8.8119	7.6229	18.2284	10.9793	11.7830	12.7084	20.8866	15.8800	19.7442
LSTM	4.4026	11.0483	4.2696	5.2706	5.6559	13.5332	6.7628	6.7668	7.4295	18.1066	10.2267	10.6434	11.0232	20.6507	15.5445	17.4355
GRU	4.3800	10.6793	4.0142	5.2468	5.5925	13.5332	5.9798	6.7075	7.3062	17.4388	9.2692	10.2745	10.0115	20.3639	14.0319	17.0238
CNN-LSTM	4.3745	10.3403	4.1279	5.1183	5.6096	13.5243	6.4372	6.5674	7.2912	18.2253	10.5793	10.5332	11.0117	20.2319	15.3219	17.4677
EMD-GRU	2.6150	5.9872	2.0438	2.5266	2.9722	7.5682	2.6702	3.7429	3.8510	9.8680	3.6771	5.5239	4.9298	11.6475	5.7202	9.2583
WPD-GRU	2.6343	4.6732	2.1107	2.0857	4.7544	9.2999	4.5272	5.3098	7.3967	15.6650	8.3896	9.1341	11.5502	21.4218	14.2254	19.1285
CEEMDAN-GRU	2.5894	5.8054	1.8043	2.4525	2.9445	7.5359	2.7131	4.1710	3.5871	9.4581	3.5030	5.3204	4.6209	11.4922	4.4362	9.1622
Proposed method	**1.3041**	**3.5929**	**1.0039**	**1.6011**	**2.3612**	**6.4616**	**2.3648**	**3.6567**	**3.3737**	**8.7188**	**3.4322**	**5.0376**	**4.3533**	**11.4485**	**4.3397**	**9.0660**

Table 4 ***NMAPE*** **values of forecasting methods in four seasons in Case II**

Method	*NMAPE* value (%)															
	10 min				30 min				1 h				2 h			
	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn	Winter
BP	3.4970	6.1587	3.2613	3.9735	4.3397	8.1291	7.4257	6.6132	5.9075	11.1271	8.1387	8.8703	9.9571	14.4128	12.4534	14.3730
LSTM	3.3403	6.4625	3.1629	3.8590	4.2738	7.6965	5.0921	5.0788	5.6157	11.1741	7.8965	8.0592	8.7830	14.5711	12.5723	12.8644
GRU	3.3492	5.9248	2.8535	3.8131	4.2516	7.7515	4.3819	5.0364	5.3553	11.1382	6.6876	7.4407	8.2593	14.2967	10.6271	11.9524
CNN-LSTM	3.3257	5.6243	3.1400	3.7543	4.2267	7.6370	4.6949	4.9520	5.4060	11.2064	8.3491	7.8989	8.7759	14.6369	11.9229	12.3634
EMD-GRU	2.0250	3.7307	1.5774	1.8841	2.2077	4.1881	2.0653	2.7372	3.0181	5.4555	2.9065	4.2004	3.9046	6.8915	4.4718	7.0097
WPD-GRU	2.2143	2.8217	1.6138	1.5849	3.7075	5.5671	3.3148	3.8462	5.5453	9.7906	6.2595	5.7264	9.1667	14.2830	10.8970	13.6730
CEEMDAN-GRU	1.9892	3.5432	1.4269	1.8137	2.3599	4.3260	2.1587	2.8888	2.7915	5.5011	2.7062	3.9539	3.6256	6.6108	3.5997	6.7228
Proposed method	**1.0576**	**2.1075**	**0.8077**	**1.1310**	**1.6262**	**3.5582**	**1.7612**	**2.3936**	**2.5549**	**5.0467**	**2.5413**	**3.6737**	**3.3377**	**6.5611**	**3.3890**	**6.7197**

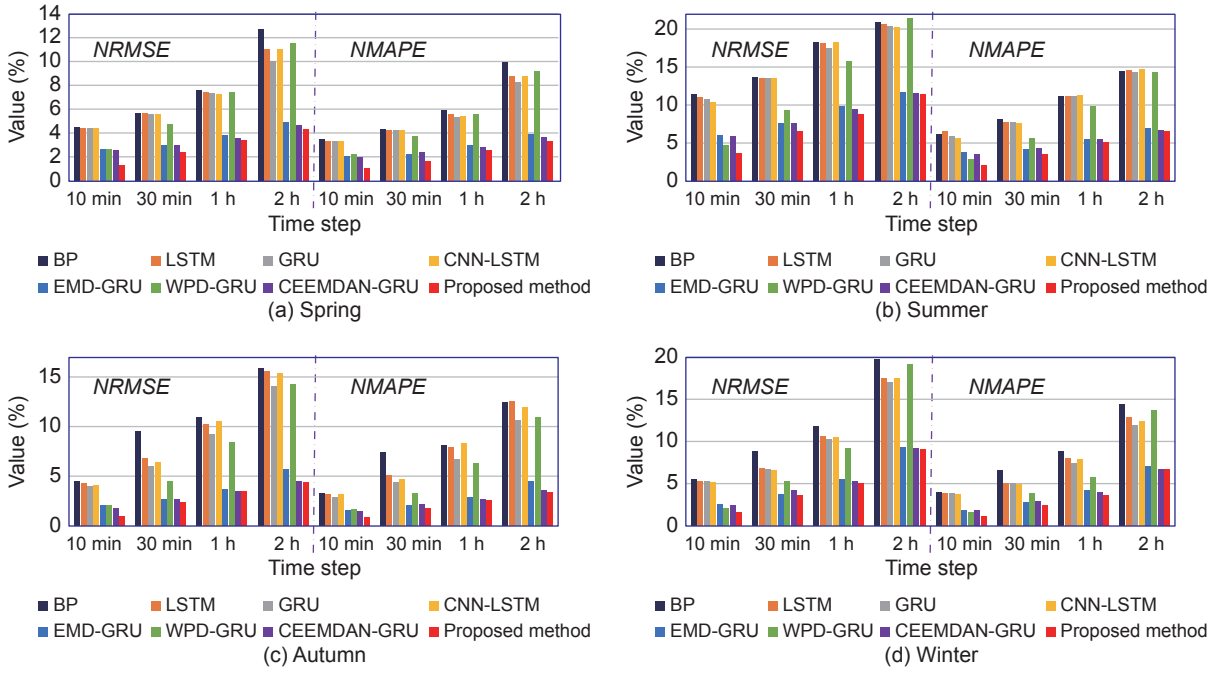

Fig. 6 Histogram of all the metrics in four seasons in Case II.

3.3377%), respectively. Similarly, in Fig. 6b (summer), compared with the simple GRU method, the *NRMSE* values decrease 66.36% and 43.78% for 10 min forecast horizon (from 10.6793% to 3.5929%) and 2 h

forecast horizon (from 20.3639% to 11.4485%), respectively, and the *NMAPE* values decrease 64.43% and 54.11% for 10 min forecast horizon (from 5.9248% to 2.1075%) and 2 h forecast horizon (from 14.2967% to 6.5611%), respectively. These results indicate that the proposed method can well adapt to complex environmental changes in different seasons and mine more realistic features from highly random WSTS.

To intuitively show the performance of the proposed method, the corresponding prediction curve at 10 min forecast horizon in spring is plotted in Fig. 7, revealing the proposed method learned the variations of real wind speed.

3.3 Model structure and parameter analysis

In order to know the reason why a high accuracy for wind speed forecasting can be achieved through the proposed method, analyzing their GRU networks structure for each component signal obtained by the time series decomposition algorithm is essential. It is necessary to properly select the hidden layer neurons number and the iteration number for prediction tasks. But there is no theoretical method for solving the problem at present. Consequently, a trial-and-error means was put to use, and thus determining these structural parameters. Taking the *AA* component for 10 min forecast horizon in spring in Case I as an example, the hidden layer neurons number was trialed from 2 to 40 in steps of 2. The iteration number was trialed from {10, 20, 30, 40, 50}. In addition, the *NRMSE* value is chosen to evaluate the accuracy of predicted values for *AA* component. The change of these two parameters on *NRMSE* value is shown in Fig. 8, from which it can be found that when the hidden layer neurons number and iteration number are 34 and 10, respectively, the metric *NRMSE* reaches its lowest

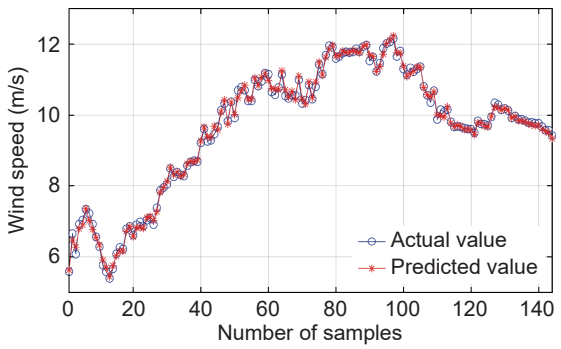

Fig. 7 Prediction curve of the proposed method at 10 min forecast horizon in spring in Case II.

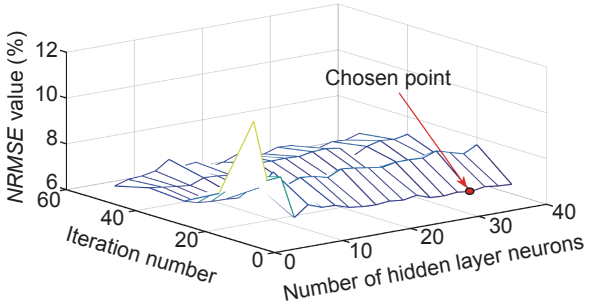

Fig. 8 Change of the hidden layer neurons number and iteration number on *NRMSE* value.

value (marked by a red circle). Similarly, all the structural parameters for Cases I and II can be obtained following this method.

4 Conclusion

With the objective of predicting short-time wind speed, a hybrid deep learning model was put forward in consideration of the stochastic and volatile nature of the processes. The model comprises two parts: time series decomposition algorithm and GRU networks. The time series decomposition algorithm realizes meritorious service in dealing with the volatility of WSTS, i.e., the complex WSTS was converted into refined sub-series via CEEMDAN and WPD. Then, for each of the resulting component signals, the GRU-based prediction model is built and trained. This stage takes the advantage of GRU network for achieving high prediction performance. Two case simulations considering four different seasons and forecast horizons were implemented, which involved two types of datasets: offshore and onshore wind farms. Comparisons with seven existing methods, namely, BP, LSTM, GRU, CNN-LSTM, EMD-GRU, WPD-GRU, and CEEMDAN-GRU, were conducted, which revealed the proposed method carried out best, pointing at the two commonly used metrics: NRMSE and NMAPE. In future work, potential influencing factors affecting its predictive ability will be investigated. It would be interesting to take advantage of different methods to enhance the predictive faculty for multiple season wind speed of the year.

Acknowledgment

This work was supported in part by the National Key Research and Development Project of China (No. 2019YFE0105300), the National Natural Science Foundation of China (No. 61972443), and the Hunan

Provincial Key Research and Development Project of China (No. 2022WK2006).

References

[1] T. Hu, W. Wu, Q. Guo, H. Sun, L. Shi, and X. Shen, Very short-term spatial and temporal wind power forecasting: A deep learning approach, *CSEE Journal of Power and Energy Systems*, vol. 6, no. 2, pp. 434–443, 2020.

[2] Y. Ren, P. N. Suganthan, and N. Srikanth, A novel empirical mode decomposition with support vector regression for wind speed forecasting, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1793–1798, 2016.

[3] Y. Wang, D. Wang, and Y. Tang, Clustered hybrid wind power prediction model based on ARMA, PSO-SVM, and Clustering Methods, *IEEE Access*, vol. 8, pp. 17071–17079, 2020.

[4] O. Karakuş, E. E. Kuruoğlu, and M. A. Altınkaya, One-day ahead wind speed/power prediction based on polynomial autoregressive model, *IET Renewable Power Generation*, vol. 11, no. 11, pp. 1430–1439, 2017.

[5] H. L. Wei, Boosting wavelet neural networks using evolutionary algorithms for short-term wind speed time series forecasting, presented at the International Work-Conference on Artificial Neural Networks, Gran Canaria, Spain, 2019.

[6] J. Bedi and D. Toshniwal, Empirical mode decomposition based deep learning for electricity demand forecasting, *IEEE Access*, vol. 6, pp. 49144–49156, 2018.

[7] M. Khodayar and J. H. Wang, Spatio-temporal graph deep neural network for short-term wind speed forecasting, *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 670–681, 2019.

[8] L. Yang, M. He, J. Zhang, and V. Vittal, Support-vector-machine-enhanced Markov model for short-term wind power forecast, *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 791–799, 2015.

[9] J. Yan, K. Li, E. -W. Bai, J. Deng, and A. M. Foley, Hybrid probabilistic wind power forecasting using temporally local Gaussian process, *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 87–95, 2016.

[10] X. Li, S. Cao, L. Gao, and L. Wen, A threshold-control generative adversarial network method for intelligent fault diagnosis, *Complex System Modeling and Simulation*, vol. 1, no. 1, pp. 55–64, 2021.

[11] M. Khodayar, J. Wang, and M. Manthouri, Interval deep generative neural network for wind speed forecasting, *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3974–3989, 2019.

[12] Q. Zhang, F. Li, F. Long, and Q. Ling, Vehicle emission forecasting based on wavelet transform and long short-term memory network, *IEEE Access*, vol. 6, pp. 56984–56994, 2018.

[13] W. Yao, P. Huang, and Z. Jia, Multidimensional LSTM networks to predict wind speed, presented at 37th Chinese Control Conference, Wuhan, China, 2018.

[14] Y. Long, T. Na, and S. Mukhopadhyay, ReRAM-based processing-in-memory architecture for recurrent neural network acceleration, *IEEE Transactions on Very Large Scale Integration* (*VLSI*) *Systems*, vol. 26, no. 12, pp. 2781–2794, 2018.

[15] C. Li, G. Tang, X. Xue, A. Saeed, and X. Hu, Short-term wind speed interval prediction based on ensemble GRU model, *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1370–1380, 2020.

[16] H. Liu, H. Q. Tian, D. F. Pan, and Y. F. Li, Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks, *Applied Energy*, vol. 107, pp. 191–208, 2013.

[17] H. Liu, C. Chen, H. Q. Tian, and Y. F. Li, A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks, *Renewable Energy*, vol. 48, pp. 545–556, 2012.

[18] J. Shi, Z. Ding, W. J. Lee, Y. Yang, Y. Liu, and M. Zhang, Hybrid forecasting model for very-short term wind power forecasting based on grey relational analysis and wind speed distribution features, *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 521–526, 2014.

[19] M. Khodayar, O. Kaynak, and M. E. Khodayar, Rough deep neural architecture for short-term wind speed forecasting, *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2770–2779, 2017.

[20] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, EMD-based temporal and spectral features for the classification of EEG signals using supervised learning, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 28–35, 2016.

[21] T. Wang, T. Liu, K. Liu, J. Jiang, L. Yu, M. Xue, and Y. Meng, An EMD-based filtering algorithm for the fiber-optic SPR sensor, *IEEE Photonics Journal*, vol. 8, no. 3, pp. 1–8, 2016.

[22] H. Liu, X. W. Mi, and Y. F. Li, Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network, *Energy Conversion and Management*, vol. 156, pp. 498–514, 2018.

[23] R. Ye, Z. Guo, R. Liu, and J. Liu, Short-term wind speed forecasting method based on wavelet packet decomposition and improved Elman neural network, presented at 2016 International Conference on Probabilistic Methods Applied to Power Systems, Beijing, China, 2016.

[24] L. Zhao, Y. Jia, and Y. Xie, Robust transcale decentralised estimation fusion for multisensor systems based on wavelet packet decomposition, *IET Control Theory and Applications*, vol. 8, no. 8, pp. 585–597, 2014.

[25] L. Han, R. Zhang, X. Wang, A. Bao, and H. Jing, Multi-step wind power forecast based on VMD-LSTM, *IET Renewable Power Generation*, vol. 13, no. 10, pp. 1690–1700, 2019.

[26] X. W. Zheng, Y. Y. Tang, and J. T. Zhou, A framework of adaptive multiscale wavelet decomposition for signals

on undirected graphs, *IEEE Transactions on Signal Processing*, vol. 67, no. 7, pp. 1696–1711, 2019.

[27] A. U. Haque, P. Mandal, J. Meng, A. K. Srivastava, T. -L. Tseng, and T. Senjyu, A novel hybrid approach based on wavelet transform and fuzzy ARTMAP networks for predicting wind farm power production, *IEEE Transactions on Industry Applications*, vol. 49, no. 5, pp. 2253–2261, 2013.

[28] X. L. Wang and H. Li, One-month ahead prediction of wind speed and output power based on EMD and LSSVM, in *Proc. 2009 International Conference on Energy and Environment Technology*, Guilin, China, 2009, pp. 439–442.

[29] M. A. Colominas, G. Schlotthauer, and M. E. Torres, Improved complete ensemble EMD: A suitable tool for biomedical signal processing, *Biomedical Signal Processing and Control*, vol. 14, pp. 19–29, 2014.

[30] A. Komaty, A. O. Boudraa, J. P. Nolan, and D. Dare, On the behavior of EMD and MEMD in presence of symmetric α-stable noise, *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 818–822, 2015.

[31] A. H. Liao, C. C. Shen, and P. C. Li, Potential contrast improvement in ultrasound pulse inversion imaging using EMD and EEMD, *IEEE Transactions on Ultrasonics*, *Ferroelectrics*, *and Frequency Control*, vol. 57, no. 2, pp. 317–326, 2010.

[32] X. J. Liu, Z. Q. Mi, P. Li, and H. W. Mei, Study on the multi-step forecasting for wind speed based on EMD, in *Proc. International Conference on Sustainable Power Generation and Supply*, Nanjing, China, 2009, pp. 1–5.

[33] P. Mandal, K. Tank, T. Mondal, C. -H. Chen, and M. J. Deen, Predictive walking-age health analyzer, *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 363–374, 2018.

[34] R. Abdelkader, A. Kaddour, A. Bendiabdellah, and Z. Derouiche, Rolling bearing fault diagnosis based on an improved denoising method using the complete ensemble empirical mode decomposition and the optimized thresholding operation, *IEEE Sensors Journal*, vol. 18, no. 17, pp. 7166–7172, 2018.

[35] A. Humeau-Heurtier, G. Mahé, and P. Abraham, Multi-dimensional complete ensemble empirical mode decomposition with adaptive noise applied to laser speckle contrast images, *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 2103–2117, 2015.

[36] S. Saroha and S. K. Aggarwal, Wind power forecasting using wavelet transforms and neural networks with tapped delay, *CSEE Journal of Power and Energy Systems*, vol. 4, no. 2, pp. 197–209, 2018.

[37] H. Liu, H. Q. Tian, X. F. Liang, and Y. F. Li, Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks, *Applied Energy*, vol. 157, pp. 183–194, 2015.

[38] W. Li, T. Logenthiran, and W. L. Woo, Multi-GRU prediction system for electricity generation's planning and operation, *IET Generation Transmission and Distribution*, vol. 13, no. 9, pp. 1630–1637, 2019.

[39] X. Luo, J. Sun, L. Wang, W. Wang, W. Zhao, J. Wu, J. -H. Wang, and Z. Zhang, Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy, *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4963–4971, 2018.

[40] National Renewable Energy Laboratory, https://www.nrel.gov, 2015.

Zhaohua Liu received the MSc degree in computer science and engineering, and the PhD degree in automatic control and electrical engineering from Hunan University, China, in 2010 and 2012, respectively. He worked as a visiting researcher in the Department of Automatic Control and Systems Engineering at the University of Sheffield, UK, from 2015 to 2016.

He is currently an associate professor with the School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan, China. His current research interests include artificial intelligence and machine learning algorithm design, parameter estimation and control of permanent-magnet synchronous machine drives, and condition monitoring and fault diagnosis for electric power equipment.

He has published a monograph in the field of biological immune system inspired hybrid intelligent algorithm and its applications, and published more than 30 research papers in refereed journals and conferences, including IEEE TRANSACTIONS/JOURNAL/MAGAZINE. He is a regular reviewer for several international journals and conferences.

Changtong Wang received the BEng degree in automation from Hunan University of Science and Technology, Xiangtan, China, in 2019. He is currently pursuing the MSc degree in automatic control and electrical engineering, at Hunan University of Science and Technology, Xiangtan, China. His current research interests include deep learning algorithm design and wind power forecasting and dispatching.

Hualiang Wei received the PhD degree in automatic control from University of Sheffield, Sheffield, UK, in 2004.

He is currently a senior lecturer with the Department of Automatic Control and Systems Engineering, the University of Sheffield, Sheffield, UK. His research focuses on evolutionary algorithms, identification and modelling for complex nonlinear systems, applications and developments of signal processing, and system identification and data modelling to control engineering.

Lei Chen received the MS degree in computer science and engineering, and the PhD degree in automatic control and electrical engineering from Hunan University, China, in 2012 and 2017, respectively.

He is currently a lecturer with the School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan, China. His current research interests include deep learning, network representation learning, information security of industrial control system, and big data analysis.

Hongqiang Zhang received the BS, MS, and PhD degrees in control science from Hunan University of Science and Technology in 2001, Hunan University (HNU) in 2004, and HNU in 2016, respectively.

He is currently a lecturer with the School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan, China. His research interests are swarm robotics system, swarm intelligence, optimization, and intelligent control.